

Penerapan Klasterisasi Menggunakan *K-Means* untuk Menentukan Tingkat Kesehatan Bayi dan Balita di Kabupaten Bengkulu Utara

Daniel Tunggono Saputro¹, Wida Pesah Sucihermayanti²

¹Program Studi Teknik Informatika, ²Program Studi Sistem Informasi,

Fakultas Teknik dan Informatika, Universitas AKI Semarang

Jl Imam Bonjol No.15 - 17, Dadapsari, Kec. Semarang Utara, Semarang, 50173, Jawa Tengah, Indonesia

Email: ¹daniel.tunggono@student.unaki.ac.id, ²222180009@unaki.ac.id

Abstract. *The health level in Indonesia is still a challenge in every central and district area, especially the health level of infants and toddlers. The Health Office in North Bengkulu Regency always strives to provide the best health services to the people in North Bengkulu. Focusing on the health level of infants and toddlers in Bengkulu Regency, this study applies clustering to help the Health Office determining their health level in each village/district. K-means algorithm is used to cluster each sub-district based on indicators of infant mortality, under-five mortality, morbidity, and nutritional status. The result showed that the processing of existing indicators is grouped into three clusters covering high, medium and low health levels.*

Keywords: *Clustering, K-means, Health, RapidMiner*

Abstrak. *Tingkat kesehatan di Indonesia masih menjadi tantangan di setiap daerah pusat maupun kabupaten, terutama tingkat kesehatan pada bayi dan balita. Dinas Kesehatan di Kabupaten Bengkulu Utara selalu berusaha untuk memberikan pelayanan kesehatan terbaik terhadap masyarakat di Bengkulu Utara. Dalam rangka memperhatikan tingkat kesehatan bayi dan balita pada Kabupaten Bengkulu, penelitian ini menerapkan klasterisasi untuk membantu Dinas Kesehatan mengetahui tingkat kesehatan bayi dan balita di setiap desa/kecamatan. Algoritma K-means digunakan untuk mengklasterisasi setiap kecamatan berdasarkan indikator angka kematian bayi, angka kematian balita, angka kesakitan, dan status gizi. Hasil proses pengolahan indikator yang ada dikelompokkan ialah tiga klaster meliputi tingkat kesehatan tinggi, sedang dan rendah.*

Kata Kunci: *Klasterisasi, K-means, Kesehatan, RapidMiner*

1. Pendahuluan

Tingkat kesehatan masyarakat di Indonesia masih menjadi tantangan yang cukup tinggi baik di tingkat pusat maupun ditingkat kabupaten terlebih terhadap kesehatan bayi dan balita. Dalam hal ini, kesehatan bayi dan balita dapat ditentukan oleh beberapa indikator yaitu Angka Kematian Bayi (AKB), angka kematian balita, angka kesakitan, dan status gizi [1]. Dari indikator tersebut dapat mengetahui dan menentukan tingkat kesehatan bayi dan balita, sehingga dapat dijadikan acuan pemerintah daerah atau tenaga kesehatan untuk lebih meningkatkan pelayanan kepada masyarakat yang membutuhkan. Hal tersebut dapat mengurangi angka kematian bayi dan balita. Selain itu, tingkat kesehatan juga dapat dijadikan acuan untuk menentukan alokasi anggaran dana dari pusat ke daerah [2].

Dinas Kesehatan Kabupaten Bengkulu Utara setiap tahunnya melakukan pendataan kesehatan penduduk termasuk bayi dan balita. Kegiatan ini dapat memberikan informasi berkaitan dengan kesehatan penduduk setiap kecamatan di Kabupaten Bengkulu Utara. Tetapi, Dinas Kesehatan belum melakukan klasterisasi pengolahan data yang berakibat tingkat kesehatan bayi dan balita di setiap kecamatan belum diketahui secara pasti. Data yang didapatkan oleh Dinas Kesehatan Bengkulu Utara hanya untuk mengolah data kesehatan laporan bulanan menjadi data laporan tahunan.

Dari permasalahan yang ada, maka diperlukan pengolahan data menjadi sebuah klasterisasi agar mempermudah pengelompokan tingkat kesehatan bayi dan balita di Kabupaten Bengkulu Utara. Penelitian yang dilakukan oleh Chandra M. mempunyai kesimpulan bahwa algoritma *K-means* pada pengelompokan balita yang memiliki gizi buruk menurut provinsi dapat diterapkan [3]. Dari hasil pengelompokan, diperoleh dua *cluster* yaitu tinggi dan rendah, terdiri dari 15 provinsi *cluster* tinggi dan 19 provinsi *cluster* rendah. Pengujian data pada *RapidMiner* menggunakan algoritma *K-means* menampilkan keakuratan data antara perhitungan manual dan sistem [3]. Penelitian ini menggunakan algoritma *K-means* sebagai metode klasterisasi pemecahan masalah dan *RapidMiner* untuk melakukan penghitungannya. Data yang digunakan adalah data mortalitas, morbiditas dan status gizi tahun 2019-2020 yang ada di Dinas Kesehatan Bengkulu Utara.

2. Tinjauan Pustaka

Tahapan dalam penelitian ini ialah mengumpulkan data, yaitu dimulai dari melakukan studi lapangan, kemudian tahap mengidentifikasi masalah untuk menentukan tujuan penelitian, kemudian tahap selanjutnya yaitu mengumpulkan, tahap selanjutnya yang dilakukan adalah rekap data untuk dilakukannya analisis data dilakukan pengoperasian algoritma *K-means* dari metode klasterisasi. Windha Mega dalam penelitiannya yang berjudul *Clustering Menggunakan Metode K-Means untuk Menentukan Status Gizi Balita* menyimpulkan bahwa algoritma *K-Means* hanya memiliki nilai akurasi 34% benar [4]. Metode yang sama akan diterapkan dalam penelitian ini, namun dengan data yang lebih banyak. Penelitian lain yang berjudul *Sistem Informasi Geografis Pemetaan Faktor-Faktor Yang Mempengaruhi Angka Kematian Ibu (AKI) Dan Angka Kematian Bayi (AKB) dengan Metode K-Means Clustering (Studi Kasus: Provinsi Bengkulu)* mempunyai kesimpulan secara keseluruhan bahwa tingkat AKI/AKB di setiap kota/kabupaten Bengkulu masih belum memuaskan, yaitu < 15% untuk AKI dan < 32,5% untuk AKB [5].

Hasil penelitian “Penerapan Data Mining untuk Pengelompokan Data Rekam Medis Pasien Berdasarkan Jenis Penyakit Dengan Algoritma *Clustering* (Studi Kasus : Poli Klinik PT. Inecda)” mempunyai kesimpulan *data mining* dengan metode Algoritma *clustering K-means* membantu pengelompokan data rekam medis pasien Poli Klinik PT. Inecda berdasarkan wilayah, jenis kelamin, dan umur. Jumlah penyakit dengan pasien yang terbanyak adalah pertama ISPA dengan jumlah pasien 1985 pasien, yang disebabkan lingkungan perumahan PT. Inecda yang merupakan perkebunan kelapa sawit dan pabrik kelapa sawit. [6]. Penelitian oleh Eni dan Siti yang berjudul *Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita* menunjukkan nilai gizi balita di Desa Sukamantri Bogor 30% balita obesitas serta 11% balita kekurangan gizi sehingga perlu adanya pendampingan dari posyandu serta puskesmas terkait kepada orang tua balita sehingga jumlah balita yang kekurangan gizi dapat menurun di tahun berikutnya [7].

Hasil penelitian Aries Aprilia berjudul *Penentuan Kategori Status Gizi Balita Menggunakan Penggabungan Metode Klasterisasi Agglomerative dan K-Means* memiliki kesimpulan jumlah klaster yang optimal untuk mengklasterkan status gizi balita sebanyak 3 klaster dengan nilai SSE 44,42754 dan jarak antar klaster 0,74069 [8]. Penelitian dengan judul *Implementasi Algoritma K-Means Clustering untuk Mengelompokkan Data Gizi Balita pada Kecamatan Garoga Tapanuli Utara* yang dilakukan oleh Dody Sitohang dan Alex Rikky menghasilkan pembagian pengelompokan data gizi balita yang belum efektif dan efisien. Hal tersebut dikarenakan belum adanya metode khusus yang mendukung keberhasilan pembagian pengelompokan data gizi balita. Proses masih dilakukan secara manual dengan melakukan pembagian kategori berdasarkan data-data berat badan dan tinggi badan yang kemudian akan dikategorikan menjadi gizi baik, sedang, dan buruk dengan nilai yang didapat pada setiap *cluster* [9].

Penelitian *Clustering Data Antropometri Balita untuk Menentukan Status Gizi Balita di Kelurahan Jumptut Rejo Sukodono Sidoarjo* oleh Amir Ali menghasilkan kesimpulan yang dapat dijadikan acuan bagi melalui bidan dan tenaga kesehatan puskesmas yang terlatih untuk melakukan sosialisasi dan pencegahan terhadap gizi buruk balita berdasarkan tinggi badan dan berat badan balita [10]. Penelitian oleh Rifki Adhitama dengan judul *Penentuan Jumlah Cluster*

Ideal SMK di Jawa Tengah dengan Metode *X-Means Clustering* dan *K-Means Clustering* mempunyai kesimpulan bahwa hasil dari kedua skenario penelitian yang telah dilakukan menunjukkan nilai DBI yang konsisten antara *X-means* dan *K-means clustering* untuk mengukur hasil *cluster* persebaran data pokok SMK dengan enam buah kriteria [11]. Berdasar tinjauan penelitian-penelitian sebelumnya, penelitian ini menerapkan Algoritma *K-means* untuk mengetahui tingkat kesehatan bayi dan balita di Kabupaten Bengkulu Utara berdasar data mortalitas, morbiditas dan status gizi. Hasil yang menunjukkan tingkat kesehatan bayi dan balita yang rendah akan menjadi perhatian khusus bagi Dinas Kesehatan Kabupaten Bengkulu Utara.

RapidMiner merupakan *platform* perangkat lunak ilmu data yang menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran penambangan teks, dan analisis prediktif. Hal ini digunakan untuk bisnis dan komersial, juga untuk penelitian, pendidikan pelatihan, *rapid prototyping* dan pengembangan aplikasi serta mendukung semua langkah dalam proses pembelajaran termasuk persiapan data, hasil visualisasi, validasi model, dan optimasi [12]. *RapidMiner* telah digunakan ribuan organisasi dan industri di seluruh dunia untuk meningkatkan pendapatan, mengurangi biaya dan mengurangi resiko dalam menjalankan organisasi tersebut. *RapidMiner* sudah dibentuk sedemikian rupa agar mudah digunakan dalam mengolah data yang akan diproses sehingga menghasilkan data hasil yang akurat sehingga dapat meningkatkan keterampilan organisasi [13].

3. Metodologi Penelitian

3.1. Wawancara dan Observasi

Wawancara dengan pihak Dinas Kesehatan Bengkulu Utara dilakukan untuk mendapatkan informasi yang diinginkan. Penulis juga melakukan pengamatan terhadap permasalahan yang diteliti untuk mendapatkan data secara tepat seperti melihat pengumpulan data bayi dan balita.

Dinas Kesehatan di Kabupaten Bengkulu Utara memiliki 22 Puskesmas yang melakukan pelayanan masyarakat di berbagai kecamatan di Kabupaten Bengkulu Utara. Dinas Kesehatan sebagai pelaku dan penanggung jawab yang diberikan pemerintah dalam menangani kasus salah satunya tingkat kesehatan bayi dan balita di Kabupaten Bengkulu Utara. Beberapa masalah terkait hal tersebut ialah: (1) Dinas Kesehatan mengumpulkan data kesehatan bayi dan balita yang didapat dari setiap puskesmas di Bengkulu Utara yang belum diproses menjadi informasi yang bermanfaat; (2) Dinas Kesehatan membutuhkan sebuah pola untuk mengetahui tingkat kesehatan bayi dan balita secara cepat dan tepat agar dapat memberikan pelayanan sesuai dengan kebutuhan.

Solusi penyelesaian permasalahan Dinas Kesehatan Bengkulu Utara mengenai tingkat kesehatan bayi dan balita ialah membuat klasterisasi tingkat kesehatan bayi dan balita di Kabupaten Bengkulu Utara. Data bayi dan balita digunakan sebagai acuan membuat klaster mengenai tingkat kesehatan bayi dan balita. Klasterisasi dapat membantu Dinas Kesehatan memberikan pelayanan kesehatan bayi dan balita secara tepat.

3.2. Data Understanding

Pembuatan klasterisasi tingkat kesehatan bayi dan balita memerlukan data acuan berupa data bayi dan balita di Kabupaten Bengkulu Utara. Indikator mortalitas, morbiditas, dan status gizi digunakan untuk menentukan tingkat kesehatan. Data diperoleh dari Dinas Kesehatan dan merupakan data pertahun sejak tahun 2019 dan 2020 dari 22 Puskesmas di Kabupaten Bengkulu Utara. Data yang dikumpulkan dan diolah mengja dikumpulkan dari. Data yang diambil penulis yaitu jumlah bayi lahir hidup, jumlah kematian bayi, jumlah balita, jumlah kematian balita, persentase status gizi kurang, persentase status gizi buruk, persentase pneumonia, persentase diare, dan persentase TBC.

Dalam pengolahannya penulis menentukan angka kematian bayi dan angka kematian balita dalam satu tahun. Begitu pula dengan persentase status gizi kurang (%) dalam waktu setahun, persentase status gizi buruk (%) dalam setahun, persentase berat bayi lahir rendah (%) dalam setahun, persentase balita pneumonia yang tercatat dalam waktu 1 tahun, persentase balita

diare yang dicatat dalam satu tahun, dan persentase balita TBC yang dicatat dalam waktu satu tahun.

Tabel 1. Data Kesehatan Bayi dan Balita Tahun 2019 dari 22 Puskesmas di Kabupaten Bengkulu Utara

Tabel 2. Data Kesehatan Bayi dan Balita Tahun 2020 dari 22 Puskesmas di Kabupaten Bengkulu Utara

Kecamatan	Bayi lahir hidup	Kemati-an bayi	Kemati-an balita	Jumlah balita	Balita gizi kurang (%)	Balita gizi Buruk (%)	BBLR (%)	Balita pende-rita Penu-moni-a	Balita pende-rita diare	Balita pende-rita TBC
Enggano	52	2	1	373	0,7	1,8	1,9	0	16,6	0
Kerkap (Lubuk Durian)	230	0	1	1309	1,6	0,7	0,9	0	6,6	125
Hulu Palik	202	0	3	1116	2,9	0,9	4,5	0	7,7	10,5
Air Napal (kerkap)	191	5	3	970	6,1	4,8	4,2	2	8,7	71,4
Air Besi (Dusun Curup)	155	1	7	935	1,2	0,4	11	10	108	62,5
TJ. Agung Palik	158	1	0	804	1,6	1,5	1,9	1	7,8	0
Arga Makmur	372	1	2	1967	1,2	1	4,3	2	23,7	96
Arga Makmur (Prumnas)	394	2	1	2071	0,5	2,4	7,1	0	30,4	77,8
Arma jaya (kemumu)	219	0	0	1136	3,9	1,5	0,5	0	23,5	7,1
Lais	228	0	0	1231	0,5	0,3	2,2	2	22,6	107,7
Air Padang	108	1	2	594	2,8	1,5	6,5	0	40,5	80
Batik Nau	223	0	1	1344	7,1	4,7	3,1	0	7,4	0
Giri Mulya (Air Bintuhan)	286	2	1	1478	1,4	0,4	5,9	5	20,1	150
Padang Jaya (air Lais)	480	5	2	3092	5,5	1	7,5	0	4,9	120
Ketahun	340	3	2	1700	0,5	0	2,9	0	3,9	111,1
Ketahun (D6 Ketahun)	386	1	1	1797	0,7	1,4	2,1	0	25,8	120
Pinang Raya (D4 Ketahun)	129	2	2	675	6,9	3,3	7	39	30	80
Napal Putih	210	0	2	1113	2,4	0,7	1	0	11,2	13,3
Ulok Kupai (Tanjung Harapan)	247	1	2	1261	5,4	1,5	2	0	12	106,3
Putri Hijau (Sebelat)	395	2	3	1848	5,8	2,6	4,3	1	21,2	119
Putri Hijau (Karang Pulau)	154	3	2	928	5,8	3,2	7,8	0	1,4	25
Marga Sakti (Suka Makmur)	259	2	1	1515	8,1	3	3,5	0	18,6	88,2

utara

Kecamatan	Bayi lahir hidup	Kemati-an bayi	Kemati-an balita	Jumlah balita	Balita gizi kurang (%)	Balita gizi Buruk (%)	BBLR (%)	Balita pende-rita Penu-moni-a	Balita pende-rita diare	Balita pende-rita TBC
Enggano	45	2	2	380	1,3	0,8	2,2	0	0	0
Kerkap (Lubuk Durian)	203	0	2	1252	3,8	0,6	1	0	5,7	0
Hulu Palik	178	0	3	1101	3,7	0,5	8,4	0	44,2	25
Air Napal (kerkap)	192	1	2	955	3,5	0,5	4,2	10,5	27,3	0
Air Besi (Dusun Curup)	178	2	7	892	1,5	1	11,8	0	76,5	0
TJ. Agung Palik	120	0	0	792	3,1	3,1	1,7	0	0	0
Arga Makmur	391	3	3	2036	1,9	1,5	3,1	4,9	14,9	0

Arga Makmur (Prumnas)	455	1	8	2166	0,4	1,3	5,9	39,2	24,6	0
Arma jaya (kemumu)	208	1	4	1196	2,6	1,2	3,8	0	6,4	0
Lais	183	0	1	1244	0,2	0,5	2,7	0	8,6	0
Air Padang	124	2	1	594	2,5	0,9	16,1	0	6	0
Batik Nau	212	0	3	1276	3,8	2,2	7,5	0	14,4	0
Giri Mulya (Air Bintuhan)	222	0	2	1430	1,8	0,7	3,2	0	2,5	0
Padang Jaya (air Lais)	476	4	9	3062	5,2	1	6,9	0	9,9	0
Ketahun	297	4	7	1539	1,2	0,4	2	0	0	0
Ketahun (D6 Ketahun)	291	4	4	1927	1,6	0,8	6,5	0	41,6	0
Pinang Raya (D4 Ketahun)	117	2	4	705	6,1	4,8	8,5	0	7,6	0
Napal Putih	188	0	1	1103	2,7	1,4	0,5	0	41,9	0
Ulok Kupai (Tanjung Harapan)	254	2	1	1222	6,2	2	6,3	0	3,4	0
Putri Hijau (Sebelat)	362	5	3	1850	2,7	1,6	4,1	2,7	13,1	0
Putri Hijau (Karang Pulau)	166	1	1	1050	6,6	2,7	7,2	0	12,4	0
Marga Sakti (Suka Makmur)	271	0	0	1487	6,9	2,1	1,1	0	14,8	0

3.3. Data Preparation

Pada tahap persiapan data ini, dimulai dengan menghitung menggunakan rumus *summary* pada excel pada Angka Kematian Bayi dan Balita dalam tiap-tiap kecamatan. Rumus yang digunakan ialah menghitung angka kematian bayi dengan indikator jumlah lahir hidup dan jumlah kematian bayi dalam satu tahun dan angka kematian balita dengan indikator jumlah kematian balita dan jumlah balita dalam satu tahun. Berikut merupakan rumus untuk menghitung angka kematian bayi:

$$IMR = \frac{d_0}{B} \times k$$

Keterangan :

d₀ = Jumlah kematian bayi yang berumur kurang satu tahun

B = Jumlah lahir hidup dalam setahun

K = konstanta (1000)

Berikut merupakan rumus untuk menghitung angka kematian balita :

$$FMR = \frac{b}{B_1} \times k$$

Keterangan :

b = Jumlah kematian balita yang berumur kurang satu tahun

B = Jumlah lahir hidup dalam setahun

K= konstanta (1000)

3.5 Normalisasi

Setelah data dipersiapkan maka dilakukan normalisasi pada data. Dalam penelitian ini semua data indikator kesehatan akan dinormalisasikan dalam tingkat 0-1. Dalam hal ini normalisasi menggunakan nilai min dan max. berikut merupakan rumus untuk normalisasi data :

$$\text{Nilai Baru} = \left(\frac{\text{nilai asal} - \text{nilai min}}{\text{Nilai max} - \text{nilai min}} \right) \times (\text{New}_{\text{max}} - \text{New min}) + \text{newmin}$$

Keterangan :

Nilai asal = merupakan nilai asli dari tiap indikator

Nilai max = nilai max dari tiap indikator

Nilai min = nilai min dari tiap indikator

Newmax = merupakan nilai tertinggi yang di tentukan dalam kasus ini (1)

Newmin = merupakan nilai terendah ditentukan. Dalam kasus ini (0)

Pengolahan data normalisasi data tahun 2019 dan 2020 menggunakan perangkat lunak *RapidMiner*. Hasil normalisasi tahun 2019 dan 2020 terlihat pada dua gambar berikut. Gambar 1 ialah hasil normalisasi data tahun 2019 di *RapidMiner* dan Gambar 2 ialah hasil normalisasi data tahun 2020 di *RapidMiner*.

Row No.	tahun	Kecam...	balita gi...	Balita gi...	BBLR(%)	balita p...	balita p...	balita p...	Angka ...	Angka...
6	2019	TJ. Agun...	0.145	0.312	0.133	0.025	0.060	0	0.165	0
7	2019	Arga Mak...	0.092	0.208	0.362	0.051	0.209	0.640	0.070	0.136
8	2019	Arga Mak...	0	0.500	0.629	0	0.272	0.519	0.132	0.054
9	2019	Arma jay...	0.447	0.312	0	0	0.207	0.047	0	0
10	2019	Lais	0	0.062	0.162	0.051	0.199	0.718	0	0
11	2019	Air Pada...	0.303	0.312	0.571	0	0.367	0.533	0.241	0.450
12	2019	Batik Nau	0.868	0.979	0.248	0	0.056	0	0	0.099
13	2019	Giri Muly...	0.118	0.083	0.514	0.128	0.175	1	0.182	0.090
14	2019	Padang ...	0.658	0.208	0.667	0	0.033	0.800	0.271	0.086
15	2019	Ketahun	0	0	0.229	0	0.023	0.741	0.229	0.157
16	2019	Ketahun ...	0.026	0.292	0.152	0	0.229	0.800	0.067	0.074
17	2019	Pinang ...	0.842	0.688	0.619	1	0.268	0.533	0.403	0.396
18	2019	Napal P...	0.250	0.146	0.048	0	0.092	0.089	0	0.240
19	2019	Ulok Kup...	0.645	0.312	0.143	0	0.099	0.709	0.105	0.212
20	2019	Putri Hija...	0.697	0.542	0.362	0.026	0.186	0.793	0.132	0.217
21	2019	Putri Hija...	0.697	0.667	0.695	0	0	0.167	0.506	0.288
22	2019	Marga S...	1	0.625	0.286	0	0.161	0.588	0.201	0.088

Gambar 1. Hasil Normalisasi Data Tahun 2019 di *RapidMiner*

Row No.	tahun	Kecam...	balita gi...	Balita gi...	BBLR(%)	balita p...	balita p...	balita p...	Angka ...	Angka ...
6	2020	TJ. Agun...	0.433	0.614	0.077	0	0	0	0	0
7	2020	Arga Mak...	0.254	0.250	0.167	0.125	0.195	0	0.173	0.188
8	2020	Arga Mak...	0.030	0.205	0.346	1	0.322	0	0.049	0.471
9	2020	Arma jay...	0.358	0.182	0.212	0	0.084	0	0.108	0.426
10	2020	Lais	0	0.023	0.141	0	0.112	0	0	0.102
11	2020	Air Pada...	0.343	0.114	1	0	0.078	0	0.363	0.215
12	2020	Batik Nau	0.537	0.409	0.449	0	0.188	0	0	0.300
13	2020	Giri Muly...	0.239	0.068	0.173	0	0.033	0	0	0.178
14	2020	Padang ...	0.746	0.136	0.410	0	0.129	0	0.189	0.375
15	2020	Ketahun	0.148	0	0.096	0	0	0	0.303	0.580
16	2020	Ketahun ...	0.209	0.091	0.385	0	0.544	0	0.309	0.265
17	2020	Pinang ...	0.881	1	0.513	0	0.099	0	0.385	0.723
18	2020	Napal P...	0.373	0.227	0	0	0.548	0	0	0.116
19	2020	Ulok Kup...	0.896	0.364	0.372	0	0.044	0	0.177	0.104
20	2020	Putri Hija...	0.373	0.273	0.231	0.069	0.171	0	0.311	0.207
21	2020	Putri Hija...	0.955	0.523	0.429	0	0.162	0	0.136	0.121
22	2020	Marga S...	1	0.386	0.038	0	0.193	0	0	0

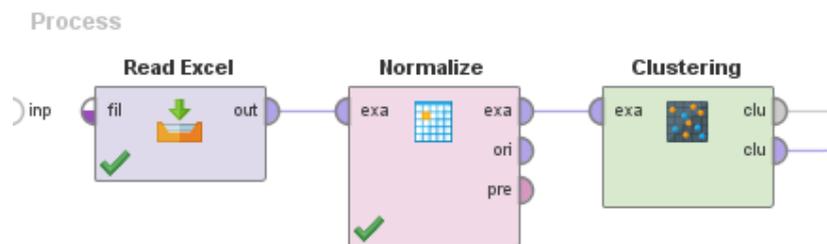
Gambar 2. Hasil Normalisasi Data Tahun 2020 di *RapidMiner*

3.6. Modeling

Setelah data siap, dilanjutkan ke proses berikutnya yaitu pemodelan. Data akan diolah menjadi sebuah *cluster* dengan menggunakan algoritma *K-means* sesuai dengan tahapan aturan algoritma *K-means*. Dari data dua tahun yang ada, masing-masing akan dilakukan proses dengan algoritma *K-means*. Setiap data akan diproses dengan urutan yang sama sehingga akan didapatkan hasil yang diinginkan. Hal yang dilakukan dalam tahapan pertama modeling yaitu memilih algoritma *cluster* yaitu *K-means*.

Langkah berikutnya adalah menentukan:

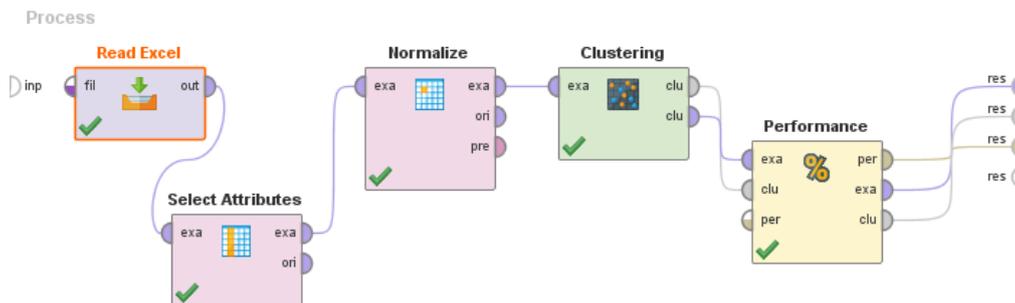
1. Menentukan jumlah kluster yang akan di bentuk. Penulis akan menetapkan 3 kluster dan hasil 3 kluster ini akan membentuk tingkat kesehatan rendah, sedang, dan tinggi. Hal itu akan dapat dilihat dari hasil prosesnya.
2. *Centroid* yang dipilih merupakan *centroid* secara acak dari ketiga kluster semuanya dipilih secara acak (*random*).
3. Menentukan kluster dengan *Euclidean distance*. Untuk menentukan jarak terdekat ke *centroid* yang dipilih sehingga mendapatkan hasil yang sesuai.



Gambar 3. Proses *Clustering* Data di *RapidMiner*.

3.7 Evaluasi

Dalam proses evaluasi *Clustering* algoritma *K-means* menggunakan *RapidMiner* maka evaluasi yang tepat adalah menggunakan *Operator Performance (Cluster Distance Performance)*. Operator ini digunakan untuk evaluasi kinerja metode *Clustering* berbasis *Centroid*. Operator ini memberikan daftar nilai kriteria kinerja berdasarkan *Centroid Cluster*.



Gambar 4. Tampilan Semua Proses Data di *RapidMiner*.

4. Hasil dan Diskusi

4.1. Hasil Pengolahan Data Tahun 2019

Gambar 5 merupakan proses *clustering* dengan syarat yang telah ditentukan. Terbentuk tiga *cluster* dengan *cluster 0* berjumlah 8, *cluster 1* berjumlah 1 dan *cluster 2* berjumlah 13. Hasil untuk titik *centroid*nya dapat dilihat pada Gambar 6.

Cluster Model

Cluster 0: 8 items
 Cluster 1: 1 items
 Cluster 2: 13 items
 Total number of items: 22

Gambar 5. Hasil Jumlah Clustering Data Tahun 2019 di RapidMiner

Attribute	cluster_0	cluster_1	cluster_2
balita gizi kurang (%)	0.768	0.092	0.144
Balita gizi Buruk (%)	0.628	0.083	0.226
BBLR(%)	0.421	1	0.258
balita penderita Penumonia	0.135	0.256	0.020
balita penderita diare	0.109	1	0.160
balita penderita TBC	0.508	0.417	0.461
Angka Kematian Bayi	0.287	0.168	0.160
Angka Kematian Balita	0.225	1	0.156

Gambar 6. Hasil Centroid Clustering di Rapidminer Data Tahun 2019

4.2. Hasil Pengolahan Data 2020

Gambar 7 merupakan proses clustering dengan syarat yang telah ditentukan. Terbentuk tiga cluster dengan cluster 0 berjumlah 7, cluster 1 berjumlah 11 dan cluster 2 berjumlah 4. Hasil titik centroid dapat dilihat pada Gambar 8.

Cluster Model

Cluster 0: 7 items
 Cluster 1: 11 items
 Cluster 2: 4 items
 Total number of items: 22

Gambar 7. Hasil Jumlah Clustering Data Tahun 2020 di RapidMiner

Attribute	cluster_0	cluster_1	cluster_2
balita gizi kurang (%)	0.778	0.270	0.317
Balita gizi Buruk (%)	0.490	0.126	0.091
BBLR(%)	0.327	0.159	0.654
balita penderita Penumonia	0	0.133	0
balita penderita diare	0.117	0.172	0.550
balita penderita TBC	0	0	0.250
Angka Kematian Bayi	0.127	0.187	0.231
Angka Kematian Balita	0.232	0.310	0.457

Gambar 8. Hasil Centroid Clustering di Rapidminer Data Tahun 2020

4.3. Kesimpulan Hasil

Dari hasil proses yang dilakukan pada kedua data yaitu data tahun 2019 dan data tahun 2020, kecamatan/desa yang memiliki indikator kesehatan tinggi merupakan kecamatan/desa yang dimana tingkat kesehatan tersebut buruk dikarenakan jumlah penderita dalam lokasi tersebut banyak. Hal ini dapat dilihat dari hasil titik pusat yang terbentuk.

Tabel 3. Hasil Clustering Data Tahun 2019

Tahun	Cluster	Angka kematian bayi	Angka kematian balita	Balita gizi kurang (%)	Balita gizi Buruk (%)	BBLR (%)	Balita penderita pnunomia	Balita penderita diare	balita penderita TBC
2019	Cluster 0	tinggi	sedang	tinggi	tinggi	sedang	sedang	rendah	tinggi
2019	Cluster 1	sedang	tinggi	rendah	rendah	tinggi	tinggi	tinggi	sedang
2019	Cluster 2	sedang	rendah	sedang	sedang	rendah	rendah	sedang	sedang

Tahun	Cluster	Angka kematian bayi	Angka kematian balita	Balita gizi kurang (%)	Balita gizi Buruk (%)	BBLR (%)	Balita penderita pnunomia	Balita penderita diare	balita penderita TBC
2020	Cluster 0	rendah	rendah	tinggi	tinggi	sedang	rendah	rendah	rendah
2020	Cluster 1	sedang	sedang	rendah	sedang	rendah	sedang	sedang	rendah
2020	Cluster 2	tinggi	tinggi	sedang	rendah	tinggi	rendah	tinggi	tinggi

Tabel 4. Hasil Clustering Data Tahun 2020

5. Kesimpulan dan Saran

Adanya klasterisasi dalam menentukan tingkat kesehatan mempermudah proses mengidentifikasi tingkat kesehatan bayi dan balita pertahun sehingga penanganan dan pemberian pelayanan akan lebih maksimal dan sesuai dengan kebutuhan.

Proses klasterisasi menggunakan algoritma *K-means* menghasilkan tiga klaster tingkat kesehatan bayi dan balita di Kabupaten Bengkulu Utara. Ketiga kluster tersebut yaitu rendah, sedang dan tinggi. Tiap tahunnya menghasilkan anggota klaster yang berbeda, dikarenakan titik pusat atau *centroid* berbeda.

Penelitian selanjutnya dapat menambahkan indikator untuk morbiditas (kesakitan) yang digunakan seperti beberapa penyakit yang sering dialami oleh bayi atau balita (contoh: demam berdarah). Selain itu media peta sebagai penggambaran hasil dari distribusi *clustering* bisa ditambahkan. Metode lain seperti algoritma *X-means* bisa digunakan sehingga didapatkan perbandingan dan hasil keputusan yang baik dan akurat.

Referensi

- [1] A. A. A. Hidayat, "Kesehatan Anak Balita di Indonesia," dalam *Pengantar Ilmu Kesehatan Anak untuk Pendidikan Kebidanan*, edisi ke 1, Jakarta, Indonesia, 2013, bab 1, hal 2.
- [2] E.A. Sari, "Penerapan Algoritma K-Means Untuk Menentukan Tingkat Kesehatan Bayi Dan Balita Pada Kabupaten Dan Kota Di Jawa Tengah", Skripsi, Jur. Teknik Informatika, Uni. Dian Nuswantoro, Semarang, Indonesia, 2015.
- [3] M. D. Chandra, E. Irawan, I. S. Saragih. A. P. Windarto, dan D. Suhendro, "Penerapan Algoritma K-Means dalam Mengelompokkan Balita yang Mengalami Gizi Buruk Menurut Provinsi", *BIOS: Jur. Teknologi Informasi dan Rekayasa Komputer*, vol. 2, no 1, hal 30-38. Mar. 2021.
- [4] W. M. P. Dhuhita, "Clustering Menggunakan Metode K-Means Untuk Menentukan Status Gizi Balita", *Jur. Informatika Darmajaya*, vol. 15 no 2, hal 160-174, Des. 2015.
- [5] K. B. Aditya, Diyah, dan Y. Setiawan, "Sistem Informasi Geografis Pemetaan Faktor-Faktor Yang Mempengaruhi Angka Kematian Ibu (Aki) dan Angka Kematian Bayi (Akb) dengan Metode K-Means Clustering (Studi Kasus: Provinsi Bengkulu), *Jur. Teknik Informatika*, vol 10 no 1, hal 59-66, Apr. 2017.

- [6] R. Ordila, R. Wahyuni, Y. Irawan dan M. Y. Sari, “Penerapan Data Mining untuk Pengelompokan Data Rekam Medis Pasien Berdasarkan Jenis Penyakit dengan Algoritma Clustering (Studi Kasus : Poli Klinik PT.Inecda)”, *Jur.l Ilmu Komputer*, vol. 9, no 2, hal 148-153, Okt. 2020.
- [7] E. Irfiani dan S.T. Rani, “Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita, *JUSTIN*, vol. 6, no 4, hal 165-172, Okt. 2018.
- [8] A. Aprilia, W. M. Rahmawati dan M. Hakimah, “Penentuan Kategori Status Gizi Balita Menggunakan Penggabungan Metode Klasterisasi *Agglomerative* Dan *K-Means*”, Seminar Nasional Sains dan Teknologi Terapan VII 2019 Institut Teknologi Adhi Tama Surabaya, hal 595-596, Sept. 2019, <https://ejurnal.itats.ac.id/sntekpan/article/view/667>.
- [9] D. W. Sitohang dan A. Rikki, “Implementasi Algoritma *K-Means Clustering* untuk Mengelompokkan Data Gizi Balita pada Kecamatan Garoga Tapanuli Utara”, *KAKIFIKOM*. Vol. 1, no 2, hal 80-92, Okt. 2019.
- [10] A. Ali, ‘Clustering Data Antropometri Balita Untuk Menentukan Status Gizi Balita Di Kelurahan Jumput Rejo Sukodono Sidoarjo”, *JATISI*, vol. 7, no. 3, hal 395-407, Des. 2020.
- [11] P. Ristoski, C. Bizer, and H. Paulheim, “Mining the Web of Linked Data with RapidMiner”, *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 35, pp. 142-151, Jul. 2015. <http://www.elsevier.com/locate/websem>.
- [12] (2021), Rapid Miner website. [online]. Tersedia: <https://rapidminer.com/why-rapidminer/> [diakses: 9 Oktober 2021].
- [13] R. Adhitama, A. Burhanuddin, R. Ananda, “Penentuan Jumlah Cluster Ideal SMK Di Jawa Tengah Dengan Metode *X-Means Clustering* Dan *K-Means Clustering*”, *JIKO*, vol. 3, no 1, hal 1-5, Apr. 2020.